

ABSTRACT OF THE DISCLOSURE

A method for processing chemical analysis data is disclosed. The method includes
5 including a step of cluster analysis, the cluster analysis using a distance metric of the
form:

$$D_{xy} = \frac{\sum_i \left(\left(\frac{x_i - c_i}{s_i} \right) - \left(\frac{y_i - c_i}{s_i} \right) \right)^2}{\sqrt{\left(\sum_i \left(\frac{x_i - c_i}{s_i} \right)^2 \right) \times \left(\sum_i \left(\frac{y_i - c_i}{s_i} \right)^2 \right)}}$$

In performance of cluster analysis, the value of the metric increases with difference in
angle α between vectors r_x and r_y starting in the co-ordinate centre and pointing at the
10 points X and Y . The value of the metric also increases with difference between lengths
of vectors r_x and r_y but this difference is normalised by their length. This means that
points located on the tail of the distribution can pass the threshold even though they are
further away from each other than points inside the standard deviation range.